## Section: Properties of Context-free Languages

Which of the following languages are CFL?

- $\mathbf{L}=\left\{a^{n} b^{n} c^{j} \mid 0<n \leq j\right\}$
- $\mathbf{L}=\left\{a^{n} b^{j} a^{n} b^{j} \mid n>0, j>0\right\}$
- $\mathbf{L}=\left\{a^{n} b^{j} a^{k} b^{p} \mid n+j \leq k+p, n>0, j>\right.$ $0, k>0, p>0\}$
- $\mathbf{L}=\left\{a^{n} b^{j} a^{j} b^{n} \mid n>0, j>0\right\}$

Pumping Lemma for Regular Language's: Let $L$ be a regular language, Then there is a constant $m$ such that $w \in L,|w| \geq m, w=x y z$ such that

- $|x y| \leq m$
- $|y| \geq 1$
- for all $i \geq 0, x y^{i} z \in L$

Pumping Lemma for CFL's Let $L$ be any infinite CFL. Then there is a constant $m$ depending only on $L$, such that for every string $w$ in $L$, with $|w| \geq m$, we may partition $w=u v x y z$ such that:
$|v x y| \leq m$, (limit on size of substring)
$|v y| \geq 1,(v$ and $y$ not both empty)
For all $i \geq 0, u v^{i} x y^{i} z \in \mathbf{L}$

- Proof: (sketch) There is a CFG G s.t. $L=L(G)$.

Consider the parse tree of a long string in L.
For any long string, some nonterminal $N$ must appear twice in the path.

Example: Consider
$L=\left\{a^{n} b^{n} c^{n}: n \geq 1\right\}$. Show $L$ is not a CFL.

- Proof: (by contradiction)

Assume L is a CFL and apply the pumping lemma.
Let $m$ be the constant in the pumping lemma and consider $w=a^{m} b^{m} c^{m}$. Note $|w| \geq m$.
Show there is no division of $w$ into $u v x y z$ such that $|v y| \geq 1,|v x y| \leq m$, and $u v^{i} x y^{i} z \in \mathbf{L}$ for $i=0,1,2, \ldots$.

Thus, there is no breakdown of $w$ into uvxyz such that $|v y| \geq 1$, $|v x y| \leq m$ and for all $i \geq 0, u v^{i} x y^{i} z$ is in $L$. Contradiction, thus, $L$ is not a CFL. Q.E.D.

Example Why would we want to recognize a language of the type $\left\{a^{n} b^{n} c^{n}: n \geq 1\right\}$ ?

Example: Consider
$L=\left\{a^{n} b^{n} c^{p}: p>n>0\right\}$. Show $L$ is not a CFL.

- Proof: Assume L is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider
$w=\ldots$ Note $|w| \geq m$.
Show there is no division of $w$ into uvxyz such that $|v y| \geq 1,|v x y| \leq m$, and $u v^{i} x y^{i} z \in \mathbf{L}$ for $i=0,1,2, \ldots$.

Example: Consider $L=\left\{a^{j} b^{k}: k=j^{2}\right\}$. Show $L$ is not a CFL.

- Proof: Assume L is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider
$w=$
Show there is no division of $w$ into $u v x y z$ such that $|v y| \geq 1,|v x y| \leq m$, and $u v^{i} x y^{i} z \in \mathbf{L}$ for $i=0,1,2, \ldots$.
Case 1: Neither $v$ nor $y$ can contain 2 or more distinct symbols. If $v$ contains $a$ 's and $b$ 's, then $u v^{2} x y^{2} z \notin \mathbf{L}$ since there will be $b$ 's before $a$ 's. Thus, $v$ and $y$ can be only $a$ 's, and $b$ 's (not mixed).

Example: Consider
$L=\left\{w \bar{w} w: w \in \Sigma^{*}\right\}, \Sigma=\{a, b\}$, where $\bar{w}$ is the string $w$ with each occurence of $a$ replaced by $b$ and each occurence of $b$ replaced by $a$. Show $L$ is not a CFL.

- Proof: Assume L is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider
$w=$ $\qquad$
Show there is no division of $w$ into $u v x y z$ such that $|v y| \geq 1,|v x y| \leq m$, and $u v^{i} x y^{i} z \in \mathbf{L}$ for $i=0,1,2, \ldots$.

Example: Consider $L=\left\{a^{n} b^{p} b^{p} a^{n}\right\} . \mathbf{L}$ is a CFL. The pumping lemma should apply!

Let $m \geq 4$ be the constant in the pumping lemma. Consider $w=a^{m} b^{m} b^{m} a^{m}$.

We can break $w$ into uvxyz, with:

Chap 8.2 Closure Properties of CFL's Theorem CFL's are closed under union, concatenation, and star-closure.

- Proof:

Given 2 CFG $G_{1}=\left(V_{1}, T_{1}, S_{1}, P_{1}\right)$ and
$G_{2}=\left(V_{2}, T_{2}, S_{2}, P_{2}\right)$

- Union:

Construct $G_{3}$ s.t. $\mathbf{L}\left(G_{3}\right)=\mathbf{L}\left(G_{1}\right)$
$\cup \mathbf{L}\left(G_{2}\right)$.
$G_{3}=\left(V_{3}, T_{3}, S_{3}, P_{3}\right)$

- Concatenation:

Construct $G_{3}$ s.t. $\mathbf{L}\left(G_{3}\right)=$
$\mathbf{L}\left(G_{1}\right) \circ \mathbf{L}\left(G_{2}\right)$.
$G_{3}=\left(V_{3}, T_{3}, S_{3}, P_{3}\right)$

- Star-Closure

Construct $G_{3}$ s.t. $\mathbf{L}\left(G_{3}\right)=\mathbf{L}\left(G_{1}\right)^{*}$
$G_{3}=\left(V_{3}, T_{3}, S_{3}, P_{3}\right)$

# Theorem CFL's are NOT closed under intersection and complementation. 

- Proof:
- Intersection:
- Complementation:

Theorem: CFL's are closed under regular intersection. If $L_{1}$ is CFL and $L_{2}$ is regular, then $L_{1} \cap L_{2}$ is CFL.

- Proof: (sketch) We take a NPDA for $L_{1}$ and a DFA for $L_{2}$ and construct a NPDA for $L_{1} \cap L_{2}$.
$M_{1}=\left(Q_{1}, \Sigma, \Gamma, \delta_{1}, q_{0}, z, F_{1}\right)$ is an
NPDA such that $\mathbf{L}\left(M_{1}\right)=L_{1}$.
$M_{2}=\left(Q_{2}, \Sigma, \delta_{2}, q_{0}^{\prime}, F_{2}\right)$ is a DFA such that $\mathrm{L}\left(M_{2}\right)=L_{2}$.

Example of replacing arcs (NOT a Proof!):

## We must formally define $\delta_{3}$. If

then

Must show

## if and only if

## Questions about CFL:

## 1. Decide if CFL is empty?

## 2. Decide if CFL is infinite?

Example: Consider
$L=\left\{a^{2 n} b^{2 m} c^{n} d^{m}: n, m \geq 0\right\}$. Show $L$ is not a CFL.

- Proof: Assume L is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider
$w=a^{2 m} b^{2 m} c^{m} d^{m}$.
Show there is no division of $w$ into $u v x y z$ such that $|v y| \geq 1,|v x y| \leq m$, and $u v^{i} x y^{i} z \in \mathbf{L}$ for $i=0,1,2, \ldots$.
Case 1: Neither $v$ nor $y$ can contain
2 or more distinct symbols. If $v$ contains $a$ 's and $b$ 's, then $u v^{2} x y^{2} z \notin \mathbf{L}$ since there will be $b$ 's before $a$ 's. Thus, $v$ and $y$ can be only $a$ 's, $b$ 's, $c^{\prime}$ s, or $d$ 's (not mixed).
Case 2: $v=a^{t_{1}}$, then $y=a^{t_{2}}$ or $b^{t_{3}}$
( $|v x y| \leq m$ )
If $y=a^{t_{2}}$, then
$u v^{2} x y^{2} z=a^{2 m+t_{1}+t_{2}} b^{2 m} c^{m} d^{m} \notin L$ since $t_{1}+t_{2}>0$, the number of $a$ 's is not twice the number of $c$ 's.
If $y=b^{t_{3}}$, then
$u v^{2} x y^{2} z=a^{2 m+t_{1}} b^{2 m+t_{3}} c^{m} d^{m} \notin L$ since $t_{1}+t_{3}>0$, either the number of $a$ 's (denoted $\mathbf{n}(a)$ ) is not twice $\mathbf{n}(c)$ or $\mathbf{n}(b)$ is not twice $\mathbf{n}(d)$.
Case 3: $v=b^{t_{1}}$, then $y=b^{t_{2}}$ or $c^{t_{3}}$
If $y=b^{t_{2}}$, then
$u v^{2} x y^{2} z=a^{2 m} b^{2 m+t_{1}+t_{2}} c^{m} d^{m} \notin L$ since $t_{1}+t_{2}>0, \mathbf{n}(b)>2 * \mathbf{n}(d)$.
If $y=c^{t_{3}}$, then
$u v^{2} x y^{2} z=a^{2 m} b^{2 m+t_{1}} c^{m+t_{3}} d^{m} \notin L$ since $t_{1}+t_{3}>0$, either $\mathbf{n}(b)>2 * \mathbf{n}(d)$ or
$2 * \mathbf{n}(c)>\mathbf{n}(a)$.
Case 4: $v=c^{t_{1}}$, then $y=c^{t_{2}}$ or $d^{t_{3}}$
If $y=c^{t_{2}}$, then
$u v^{2} x y^{2} z=a^{2 m} b^{2 m} c^{m+t_{1}+t_{2}} d^{m} \notin L$ since $t_{1}+t_{2}>0,2 * \mathbf{n}(c)>\mathbf{n}(a)$.
If $y=d^{t_{3}}$, then
$u v^{2} x y^{2} z=a^{2 m} b^{2 m} c^{m+t_{1}} d^{m+t_{3}} \notin L$ since $t_{1}+t_{3}>0$, either $2 * \mathbf{n}(c)>\mathbf{n}(a)$ or
$2 * \mathbf{n}(d)>\mathbf{n}(b)$.
Case 5: $v=d^{t_{1}}$, then $y=d^{t_{2}}$
then $u v^{2} x y^{2} z=a^{2 m} b^{2 m} c^{m} d^{m+t_{1}+t_{2}} \notin L$ since $t_{1}+t_{2}>0,2 * \mathbf{n}(d)>\mathbf{n}(c)$.
Thus, there is no breakdown of $w$ into uvxyz such that $|v y| \geq 1$,
$|v x y| \leq m$ and for all $i \geq 0, u v^{i} x y^{i} z$ is in $L$. Contradiction, thus, $L$ is not a CFL. Q.E.D.

